

Attachment A

09/17/553



Blast 2 Sequences results

PubMed | EBI | BLAST | GMIM | Taxonomy | Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTN 2.1.2 (Nov-13-2000)

Match: 5 Mismatch: 2 gap open: 1 gap extension: 2
x dropoff: 100 expect: 10.000 wordsize: 7 Filter Align

Sequence 1 of 4538106 Sequence 3 from Patent WO9740167. Length 8209 (1 .. 8209)

Sequence 3 gi_9630311 Gibbon leukemia virus, complete genome Length 8088 (1 .. 8088)

2

1

----- The statistics (bitscore and expect value) is calculated based on the size of nr database

----- If protein translation is reversed, please repeat the search with reverse strand of the query sequence.

Score = 1583 bits (21250), Expect = 0.0
Identities = 5776/8943 (64%), Gaps = 1625/8943 (18%)
Strand = Plus / Plus

```

Query: 2 tggtgtacgactgtg-ggcccccag--cgcgt-tg--gaataaaaatcctttgtgtttt
Sbjct: 12 tagag-a-gactgagccggggtaaccg-tgtgtccaaa--cctttgtgtat

Query: 56 gcatcaaga-ccgct--tctcgtagtgat--ttg---gggtgtcgccct--ct-tccga-
Sbjct: 67 goatoggagccg-eggtatgt---tgtttttggggggttctccaaactat-tgac

Query: 104 -ggccggac-gagggggattgt-tcttttactggcctttcatgggtcgtrgg-ccggg
Sbjct: 122 cggcc--acttcggggg--tctcacattggggc--tc---atccc--ggatccaa

Query: 160 aa---atctgcgaccacc--cctacacc--c--gag--aaccgacttrgg--ag-gta
Sbjct: 168 aaccccccacccagggaccaccgacc---caccacgggaggtaa--g-c-tggccagcg-a

Query: 205 aag--g-g-ata-cctt-tg-g---aa--chhhhhhhhhhhhhnc-ggcc--g-gcgtc-
Sbjct: 220 ccgttgtgtctcgcttcrgtgcataagtc-cgtaattctgactgtcttgtgtctc

Query: 247 ---tctgt-tctga---gt--gtctg--t---tttggtga-tgcgcgc---tttgg
Sbjct: 279 gtttgtgtctgagaccgtaaactctgactgcctt--gtaaatgcgcgcattttttgg

Query: 287 tttgcag-cgttccttcagaccgt-aagga--ctg-gag-qac-tgtga-t--c-agca
Sbjct: 337 ttt-cagtcgtt--tc-cgg---gtgaatcactctgcgagtqacgtgtgagtagcgaaca

Query: 336 gacgtgtctaggaggatcacaggtt-gccacccctgggggacgcggggggagggt---ggggg-
Sbjct: 390 gacgtgttcgg-ggctcacccgcctgttaatccaggggagacgtcccaaggaa--tcaggggag

```


Digitized by srujanika@gmail.com

Query: 1647 ttcttgcaga--ggccatggaagccttcaggcggttcacccctttgtatccccaccc
<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?1> 21/02/01

Blast Result

<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?1>

21/02/01

BLAST Result

Query: 2274 -ccat taggaaaacta---aaag nnnnnnnnctgggtgatgggtgcacagg--gca
 REVERSE TRANSCRIPTASE 38 P L G K L K E K K S W V M G A T G Q
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Sbjct: 2305 ccc a-tgggaaaactaagggtccagacggacggatcg---ga-aggagcgacaggcaggca
 pol polyprotein 38 P M G K V G S R R T V V E G A T G S

Query: 2327 acgg-cagtatccatggactaccgaagaaccgtt--gacttggga-gtgggacgggtaa
 REVERSE TRANSCRIPTASE 55 R Q Y F W T T R R T V D L G V G R V
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Sbjct: 2360 a-ggtc--tacccctggaccacaaaang-actttaaaatt-ggacat-aaacaagtga
 pol polyprotein 56 K V Y P W T T K R L L K I G H K Q V

Query: 2383 cccactcgttctggteat-ccctgagtgcccag-tacccct-tctaggtagagacttac
 REVERSE TRANSCRIPTASE 74 T H S F L V I P E C P V P L L G R D L
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Sbjct: 2414 cccactcgttctggteat-ccctgagtgcccag-tacccct-tctaggtagagacttac
 pol polyprotein 74 T H S F L V I P E C P V P L L G R D L

Query: 2440 tgaccaaga-t-gggatc aaatttc--tttt--gaac-aaggaagaccagaagtgtc-
 REVERSE TRANSCRIPTASE 93 L T K M G A Q I S F E Q G R P E V S
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Sbjct: 2471 taaccaa-actaaag-gccaga-tccatggggatcgg--ctggagg--gcccacaggtaaca
 pol polyprotein 93 L T K L K A Q I Q F S A E G P Q V T

Query: 2492 tggaaataacaaacccatcaactgtgtgacc--cccc-aa--ttagatgtaaatcga
 REVERSE TRANSCRIPTASE 111 V N N K P I T V L T L Q L D D E Y R
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Sbjct: 2524 tggggagaac--gccc-t-actatg-tg-cctggatcataacccatggaaagaataccga
 pol polyprotein 111 W G E R P T H C L V L N L E E E Y R

Query: 2547 ctat-ttcctccaaatgaaatcgatcaagatat--a-cagtcctgggtggaggcgt
 REVERSE TRANSCRIPTASE 129 L Y S P Q V K P D Q P I Q S W L R Q
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Sbjct: 2578 ctatcgaaaazgoc-agt--acccttc---tacccatggatcataacccatggaaagaataccga
 pol polyprotein 129 L H E K F V P S S T D P S W L Q Z

Query: 2602 ttcccca-aggctggcagaaacccgcaggatgggttggcaaagcaagtcccccaag
 REVERSE TRANSCRIPTASE 147 F P Q A W A E T A G M G L A K Q V P P Q
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Sbjct: 2629 ttcccaactgtatggcagaaagagccgcatggactatccaaatcaag-cccaac-cag
 pol polyprotein 146 F P T V W A E R A G M G L A N Q V P P

Query: 2661 -gttattca-a-ctgaaggcca-gtg-ctacaccgtatcag-tcagacagtacccctg
 REVERSE TRANSCRIPTASE 167 V I Q L K A S A T P V S V R Q Y P L
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Sbjct: 2687 tggtagt-agatc-aagatcaggatgcct-caccatggatgttc-gacaaatatccatg
 pol polyprotein 168 V V V E L R S G A S P V A V R Q Y P M

Query: 2715 agtagagaggctcgagaaggaaatggccatgttca-aagattaatcc-a-ac--agg
 REVERSE TRANSCRIPTASE 185 S R E A R E G I W P H V Q R L I Q Q
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Sbjct: 2743 agcaaaggaggctggaaaggatcagacccacatcccaag-ctttagacatagg
 pol polyprotein 184 S K E A R E G I R P H I Q K F L D L G

Query: 2770 gcatccatgt-tctgtccaaatc-cccttggaaatctccc-ctgtatccgggttaggaac
 REVERSE TRANSCRIPTASE 203 G I L V P V Q S P W N T P L L P V R K
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Sbjct: 2799 g-gtcttggatccatc-ccgtccccc-tggatc-acctotgtatccggaaaggc
 pol polyprotein 202 V L V P C R S P W N T P L L P V K K

Query: 2827 ctggaccaatgattatcgaccgtacaggacttggatggatcaataaagggtgcagg
 REVERSE TRANSCRIPTASE 222 P G T N D Y R P V Q D L R E V N K R V Q
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Sbjct: 2855 caggaccaatgactatggccatgttcaagacatggatggatcaataaagggtacagg
 pol polyprotein 221 P G T N D Y R P V Q D L R E I N K R V Q

Query: 2887 acataca-cccaacggtcccgaaacccttataa-acatgttggatggatccggatggatgg
 http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?1 21/02/01

Blast Results

REVERSE TRANSCRIPTASE 242 D I H P T V P N P Y N L L S A L P P E R
 Sbjct: 2916 atattccatccc-acagttcccaaacccttacaatcttc-tgagttcccttccggct-agg-
 pol polyprotein 241 D I H P T V P N P Y N L L S S L P P S

Query: 2946 ga-ac-tggtaacaaagtattgg-a-cttaaaagatgootttcttgoatgagattaa-c-
 REVERSE TRANSCRIPTASE 262 N W Y T V L D L K D A F F C L R L H
 Sbjct: 2971 tatacttggtaactcagtcttagatc-tcaaggatgccttttctgcctcaggctacatcc
 pol polyprotein 260 Y T W Y S V L D L K D A F F C L R L H P

Query: 3001 ccactagccaaaccacttttgccttcgaatggagagatccag-gracggaaagaacggg-
 REVERSE TRANSCRIPTASE 290 P T S Q P L F A F E W R D P G T C R T G
 Sbjct: 3030 caac-agccagccgttgcgttcgactggaaagacccagaaaaaggtaa-cacagg-
 pol polyprotein 280 N S Q P L F A P E W K D F E K G N T G

Query: 3060 cagttcaccttggacccgactgccccaaagggttcaagaactc-cccac-catcttgcg-
 REVERSE TRANSCRIPTASE 300 Q L T W T R L F Q G F K N S P T I F D
 Sbjct: 3098 cagctgacacttggacccggctaccacaagggttcaagaacttccc-actc-tcttcgac-
 pol polyprotein 299 Q L T W T R L P Q G F K N S P T L F D

Query: 3118 aagccctacacaggacctggc-caacttcaggatc--caacacccctcagggtgacccctcc
 REVERSE TRANSCRIPTASE 319 E A L H R D L A N F R I Q H P Q V T L
 Sbjct: 3146 aggccctccaccaggatttggctc-cctttagggccctcaac-ccc-cagggtgttac
 pol polyprotein 318 E A L H R D L A P F R A L N P Q V V L

Query: 3175 tccagtagtggatgacctgtttctggcgggagccaccaa-aa-g--gactgcttaga-
 REVERSE TRANSCRIPTASE 338 L Q Y V D D L L A G A T K Q D C L E
 Sbjct: 3203 tccaaatgtggacgacct-ct--tggtggccggcccccacatatgaagactgcaaaa-aa-
 pol polyprotein 337 L Q Y V D D L L V A A P T Y E D C K K

Query: 3232 gtac---gaaggcactactgtggatt--gtctgacacttggctcacagagccctctgtctaa
 REVERSE TRANSCRIPTASE 357 G T K A L L E L S D L G Y R A S A K
 Sbjct: 3260 gaacacagaa-gctct--tacaggatgtt-aa-taagtgggttacgggtatggctaa
 pol polyprotein 356 G T Q K L L Q E L S K L G Y R V S A K

Query: 3287 gaaggcccagatttg-caggagagggtaacata-cttgggtacagt-ttgc-ggggg-
 REVERSE TRANSCRIPTASE 376 K A Q I C R R E V T Y L G Y S L R G
 Sbjct: 3315 gaaggcccagctctgcca-gagagaactcacatc-tgggtac-ctact-caaggaa-
 pol polyprotein 375 K A Q L C Q R E V T Y L G Y L L K E

Query: 3343 gggagagatggotga-oggagggaaaggaaactg-tagtcagatccggccccaa-
 REVERSE TRANSCRIPTASE 394 G Q R W L T E A R K K T V V Q I P A P T
 Sbjct: 3371 gaaaaagatggtaacccca-gccggaaaggctactgtta-tgaaaatccctgttcc-
 pol polyprotein 393 G K R W L T P A R K A T V M K I P V P T

Query: 3401 cacagccaaacaactgtggagatgttttggggacactggattttgcagactgtggatccc
 REVERSE TRANSCRIPTASE 414 T A K Q V R E F L G T A G F C R L W I F
 Sbjct: 3429 gaccccccagacaggctcgtaatrtctaggactqcggatttgcaggcttggatccc
 pol polyprotein 413 T P R Q V R E F L G T A G F C R L W I P

Query: 3451 ggggtttgcacccatgttgc-cccactctaccgcataaccaaagaaaggggattt-
 REVERSE TRANSCRIPTASE 434 G F A T L A A P L Y P L T K E K G G F
 Sbjct: 3489 tgggtttgccttcctggctgtgcaccc-ttgttaccccttaacaaaaga----gagca-tc-
 pol polyprotein 433 G P A S L A A P L Y P L T K E S I

Query: 3520 cc-----tgggttcttgc---caccagaaggcatttgc-tataaaaaggccctq
 REVERSE TRANSCRIPTASE 453 S W A P E H Q K A F D A I K K A L

Blast Result

Subject: 3541 cctttatattgg--actgaggaaatcagcaggctttg-ccacataaaaaagcctg
 pol polyprotein 450 P F I W T E E X Q Q A F D H I K K A L

Query: 3570 ctg--agcgcacctgc-tctggccctccctgacgttaactaaaccctttaccccttatgtg
 REVERSE TRANSCRIPTASE 470 I S A P A L A L P D V T K P F T L Y V
 |||| |||| |||| | |||| |||| | |||| |||| | |||| |||| | |||| |||| |
 Subject: 3598 ctgtcagc-c-cctgcatt-tggccctcccaagacctcacaaggccattcacttatata
 pol polyprotein 469 L S A P A L A L P D L T K P F T L Y Y

Query: 3627 gatgagcgtaaag--ggagttagcccgaggagtttaacccaaacccttaggaccatggagga
 REVERSE TRANSCRIPTASE 489 D E R K G V A R G V L T Q T L G P W R
 |||| |||| | |||| |||| | |||| |||| | |||| |||| | |||| |||| |
 Subject: 3655 gatgagag--agcggcgtggccggggagtgctcactcagacttttaggaccctggcgc
 pol polyprotein 488 D E R A G V A R G V L T Q T L G P W R

Query: 3685 gacctgttgccctacctgtcaaagaagcttgatectgttagccagtgggtggcc--cgtatg
 REVERSE TRANSCRIPTASE 508 R P V A Y L S K K L D P V A S G W P V C
 |||| |||| | |||| |||| | |||| |||| | |||| |||| | |||| |||| |
 Subject: 3713 ggccaaatgtactatctataaaaaaaactggatccgggtggccaggggtggcaac--ctg
 pol polyprotein 507 R P V A Y L S K K L D P V A S G W P T C

Query: 3743 tctgaaggatctcgcagctgtggc-catactgtcaaggacgtgacaaatgtactttgg
 REVERSE TRANSCRIPTASE 528 L K A I A A V A I L V K D A D K L T L
 |||| |||| | |||| |||| | |||| |||| | |||| |||| | |||| |||| |
 Subject: 3771 cctggaaagcgggttcagcagtagcacttctcaaaagacgtgataagttaaaccttgg
 pol polyprotein 527 L K A V A A V A L L L K D A D K L T L

Query: 3802 gacagaatataactgtataatagcccccat-gcattggagaa-catcgatggcagcccc
 REVERSE TRANSCRIPTASE 547 G Q N I T V I A P H A L E N I V R Q P P
 |||| |||| | |||| |||| | |||| |||| | |||| |||| | |||| |||| |
 Subject: 3830 gacaaaatgtactgtgtattgtccatagc-ctcga-aagcatcgatgtcgccaaacccc
 pol polyprotein 546 G Q N V T V I A S H S L E S I V R Q P P

Query: 3860 agaccgtatggatgaccacgccccatgaccactataaaagctgt-ttcaca-gag
 REVERSE TRANSCRIPTASE 567 D R W M T N A R M T H Y Q S L L I T E
 |||| |||| | |||| |||| | |||| |||| | |||| |||| | |||| |||| |
 Subject: 3888 cgaccgggttatgaccatgccaatgactcattaccagagctgt-t-aaatgaa
 pol polyprotein 566 D R W M T N A R M T H Y Q S L L I N E

Query: 3918 agggta-c-ttt-cgtccaccaggccg-cctctcaaccctgccactcttcgtgaaga
 REVERSE TRANSCRIPTASE 586 R V T F A P P A A L N P A T L I P E E
 |||| |||| | |||| |||| | |||| |||| | |||| |||| | |||| |||| |
 Subject: 3946 agggt-aregtttgcgc-cc-cctgtgtc-ctaaaccctacttccagtcga
 pol polyprotein 585 R V S F A P P A V L N P A T L L P V E

Query: 3974 gactgatg--a-accagtgtactcatgatgtccatca-actat--t-gatgtggagactg
 REVERSE TRANSCRIPTASE 605 T D B F V T H D C H Q L L I B E T
 |||| |||| | |||| |||| | |||| |||| | |||| |||| | |||| |||| |
 Subject: 4002 gtctggaaaggccacccttccatgt-cacagg--tgc--tcaga-atacctcg-ccgaagaaactg
 pol polyprotein 604 S E A T P V H R C S E I L A E E T

Query: 4027 g-gtccg-caaggaccta-caga-cataccgtactggagaagtgtcaacctgg-t
 REVERSE TRANSCRIPTASE 622 G V R K D L T D I F L T G E V L T W P
 |||| |||| | |||| |||| | |||| |||| | |||| |||| | |||| |||| |
 Subject: 4055 gaact-cgacga-gacc-tagaagacca-accattggccgg-g-gtgcac3ctggat
 pol polyprotein 621 G T R R D L E D Q P L P G V P T W Y

Query: 4082 cactgacggaaggcgtatgtgg---ggaaggtaaga-ggatg-gtggggccgc-agt
 REVERSE TRANSCRIPTASE 640 T D G S S Y V V E G K R M A G A A V
 |||| |||| | |||| |||| | |||| |||| | |||| |||| | |||| |||| |
 Subject: 4108 -acagacgtgtacgag-t-t-tcatcacggaaaggtaa-acgga-gaocagggccccaa-t
 pol polyprotein 639 T D G S S F I T E G K R R A G A P I

Query: 4136 ggtggacgggaccgcacgtatggccagcagcgttgcggaaaggacttcagcgaaaa
 REVERSE TRANSCRIPTASE 659 V D G T R T I W A S S L P E G T S A Q K
 |||| |||| | |||| |||| | |||| |||| | |||| |||| | |||| |||| |
 http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?1 21/02/01

Blast Result:

Sbjct: 4161 cgttagatggcaagcggacggtatgggctagcagcctgccagaaggtagtcagccccagaa
pol polyprotein 657 V D G K R T V W A S S L P E G T S A Q K

Query: 4196 ggctgagctca-tggccctcacgcaagctttgcggctggccgaagggaaatccataaaca
REVERSE TRANSCRIPTASE 679 A E L M A L T Q A L R L A E G K S I N
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Sbjct: 4221 ggctgaact-agttagcctgacgcaggcattacgcctggccgaaggaaaaacatcaaca
pol polyprotein 677 A E L V A L T Q A L R L A E G K N I N

Query: 4255 ttatacggacagcaggatgccttgcgactgcacacgtaa-cacggggccatctataa
REVERSE TRANSCRIPTASE 698 I Y T D S R Y A F A T A X V H G A I Y K
 || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Sbjct: 4280 totacaoggacagcaggatgccttgcactgtca-tttcatgggcaataataa
pol polyprotein 696 I Y T D S R Y A F A T A X I H G A I Y K

Query: 4313 acaaagggggttgcattcacgcaggggggaaat-aaagaacaaggaggaaattctaa
REVERSE TRANSCRIPTASE 718 Q R G L L T S A G R E I K N K E E I L
 || || ||||| ||||| ||||| ||||| ||||| ||||| |||||

Sbjct: 4336 gcagagggggctgtcaacttctgtggaaaagatataaaa-aacaaagaggaaatt-ttg
pol polyprotein 716 Q R G L L T S A G K D I K N K E E I L

Query: 4378 g-cttattagaagecttacatttgcggaaaaaggotatattataactgttgcggaaat
REVERSE TRANSCRIPTASE 737 S L L E A L H L P K R L A I I H C P G H
 || || ||||| ||||| ||||| ||||| ||||| ||||| |||||

Sbjct: 4396 gcccgtctagggccatccatccctccctaggccgggtgcgcattatccactgtcctggccac
pol polyprotein 735 A L L E A I H L P R R V A I I H C P G H

Query: 4431 cagaaggccaaag-atctcatat--ctagagggaaaccagatggctgaccg-ggttgcacaa
REVERSE TRANSCRIPTASE 757 Q K A K D L I S R C N Q M A D P V A K
 || || ||||| ||||| ||||| ||||| ||||| ||||| |||||

Sbjct: 4456 cag--aggggaaagtaac-cctgtggccactggaaacccggaggggccga-cgaggctgcaaa
pol polyprotein 755 Q R G S N P V A T G N A R A D B A A K

Query: 4487 gcagggcagccaggctgttaacc---tctgcctataatagaaaegccaaagcc-cca
REVERSE TRANSCRIPTASE 776 Q A A Q A V N I L P I I E T P K A P
 || || ||||| ||||| ||||| ||||| ||||| |||||

Sbjct: 4512 gcaagccgccc---ctgtcgaccagactgtggc---aggAAC-tacaaaaccccaa
pol polyprotein 774 Q A A I S T R V L A G T T K P Q

Query: 4542 gaacccaga-cgacagtacaccctagaagactggcaagagaaaaagatagaccagg
REVERSE TRANSCRIPTASE 794 Z P R R Q Y T L E D W Q E I K K I D Q R
 || || ||||| ||||| ||||| ||||| |||||

Sbjct: 4561 g-agcca-atcg-ag---cc---g---ct---caaga---aa---agaccag--
pol polyprotein 790 E P I E P A Q E K T R

Query: 4601 ctctgagactccggaggggacccgtataccatcatatgggaaaggaaatctgccccacaa
REVERSE TRANSCRIPTASE 814 S E T P E G T C Y T S Y G K E I L P H K
 || || ||||| ||||| ||||| ||||| |||||

Sbjct: 4593 ---g----cc-ga-ggga---gct-cacc-cct---ga---ccgg--ggaa-aa
pol polyprotein 800 P R E L T P D K G K

Query: 4661 agaagggttagaatatgtccaa-cagatcacatc-gtctaacccacccatggaaacta-aaca
REVERSE TRANSCRIPTASE 834 E G L E Y V Q Q I H R L T H L G T K H
 || || ||||| ||||| ||||| ||||| |||||

Sbjct: 4623 agaa---tt---cat-t---aageggttacatcagt-taactcactttagg-accagaaaa
pol polyprotein 811 E F I K R L H Q L T H L G P E K

Query: 4718 -c--ctgcagc-agttggcag-aacatcc-ccttatacgttctgaggctaccaggagt
REVERSE TRANSCRIPTASE 852 L Q Q L V R T S P Y H V L R L P G V
 || || ||||| ||||| ||||| ||||| |||||

Sbjct: 4671 gcttcctcaacttag-tgaacccgtaccagcctctcatc---ccgaaacct-cca-a-t
pol polyprotein 826 L L Q L V N R T S L L I P N L Q

Query: 4772 ggctg-a-ctcg-ctggcacaacatt--gtgt---gc-cctgcacagctggtaatgc--t
REVERSE TRANSCRIPTASE 871 A D S V V K H C V P C Q L V N A
 || || ||||| ||||| ||||| ||||| |||||

Sbjct: 4721 --ctgcagttcgcgaagtc-accagtcaagtgcaggcttgcacca-tgactaatgcgggt
pol polyprotein http://www.ncbi.nlm.nih.gov/blast/bl2seq/blast2.cgi?1

21/02/01

卷之三

11. *Salvinia molesta* 843 S A V B E V T S Q G Q A C A M T T N A V

Query:	4876	actggaaagtggacttcactggaggtaaagcc-ggctaaatccggaaacaaatataatctatgg
REVERSE TRANSCRIPTASE	905	X W E V D F T E V K P A K X G N K Y L L
		! !
Sbjct:	4829	actggggaggtagacttcacagaaaatcaggctgg-tcggtatggaaacaatgtatctgtta
pol polyprotein	879	X W E V D F T E I K P G R Y G N K Y L L

```

Query: 4935 gcttttttagacacccctt-tcaggatggtagaggcttatccactaa---gaaa-gag
REVERSE TRANSCRIPTASE 925 V F V D T F S G W V E A Y P T K K E
|| || || || || || || || || || || || || || || || || || || || || || |
Sbjct: 4888 gtattcatagata-cttctccggatgggtagaagctttccatccaaaactggaaacg-g
pol polyprotein 699 V F I D T P S G W V E A F P T K T S T

```

```

Query: 4939 acctcaaccgtggctaaagaaaatactggaggaaattttccaagatggaaatccct
REVERSE TRANSCRIPTASE 943 T S T V V A K K I L E E I F P R F G I P
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Sbjct: 4946 ccct-aatcg-tctg--aaaaaaaaatattagaagaattctaccccgcttogggatccct
pol polyprotein 918 A L I V C K K Y L E E I L P R F G I P

```

Query:	5049 aaggtaatagggtcagacaatggtcgg-ctttcgttgcggccaggtaagtcaaggactggc
REVERSE TRANSCRIPTASE	963 K V I G S D N G P A F V A Q V S Q G L A
Sbjct:	5002 aaggtaatagggtcagacaatggccggcctt-gttgtctcaggtaagtcaaggactggc
pol polyprotein	937 K V I G S D N G P A F V A Q V S Q G L A

Query: 5167 aggttagagaggatgaataccataaaggagacccctt-accddatc-gatccatagac
REVERSE TRANSCRIPTASE 1002 Q V E R M N R T I K E T L T K L T T E T
 ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct: 5120 aggttagaaagaatgaaacagaacaattasagaga-ccttgaccaaaattag-ccttagagac
pol polyprotein 975 Q V E R M N R T I K E T L T K L A L E T

Query: 5343 -at-ttactc-a-cttaaaggcttttagaaat-tgttaagg-acccaaatac-tgggaccaga
REVERSE TRANSCRIPTASE 1076 L F S R L K A L E W V R Q R K W K Q
 Sbjct: 5343 at-ttactc-a-cttaaaggcttttagaaat-tgttaagg-acccaaatac-tgggaccaga
pol polyprotein 1050 F T K L K A L E I V R T Q I W D Q
<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?1> 21/02/0

Blast Results

Query: 5443 tccggggaggccatctcag-gaggagac--ttgcaagttccaca-tcgcttccaaatggta
REVERSE TRANSCRIPTASE 1094 L R E A Y S G G D L Q V P H R F Q V G
 Sbjct: 5396 tcaaagagggtgt-aatgcctgg-taccgtaaacaa-tccctcaccgg-ttccagggtccgg
 pol polyprotein 1068 T K E V Y K P G T V F Y P H P F Q V G

Query: 5499 gattc-agt-ctatgttagacgccaccg-tgcaggaaacctcgagactcggtggaaaggaa
REVERSE TRANSCRIPTASE 1113 D S V Y V R R H R A G N L E T R W K G
 Sbjct: 5452 ga-tcaagtgt-tgtcagacgccatcgaccga-gcagcttgcgcctcggtggaaaggc
 pol polyprotein 1067 D Q V L V R R H R P S S L E Y R W K G

Query: 5556 ccttatactcgtaattt-ttgcaccacaccaacggctgtgaaaggatcgaggatccccacactg
REVERSE TRANSCRIPTASE 1132 P Y L V L L T T P T A V K V B G I P T W
 Sbjct: 5509 ccataacctgg-tttgtactaccggacccgcgtaaaatcgatggattgtgtgcctg
 pol polyprotein 1106 P Y L V L L T T P T A V K V D G I A A W
 env protein 1 M V L L P

Query: 5615 gatccatgcatcccacgttaagccggcgccacccctcccg-actc-g--ggg--tggagag
VIRAL ENVELOPE PROTEIN 1 H H F T L S R R H L P T R G G E
REVERSE TRANSCRIPTASE 1152 I H A S H V K P A P P P D S G W R
 Sbjct: 5568 ggtccatgtcttcacccaaacccgtgcaccacccctcggtac-cagatgagtgcctgg-gag
 env protein 6 G S M L L T S N L H H L R H Q M S P G S
 pol polyprotein 1126 V H A S H L K P A P P S A P D E S W E

Query: 6668 c-cgaaaagactgagaatcccccttaagcttcgcctcca-tc-gc--ctgggtcccttac--
VIRAL ENVELOPE PROTEIN 17 P K R L R I P L S F A S I A W F L T
REVERSE TRANSCRIPTASE 1169 A E K T E N P L K L R I H R L V P Y
 Sbjct: 5626 ctggaaaagactgataccctttaagct--gcgt--attcggccggc-gg--cggggacga
 env protein 25 W K R L I I L L S C V F G G G G T
 pol polyprotein 1145 L E K T D H P L K L R I R R R K R D E
 P

Query: 5721 -tct--aacaat---aactcc--ccaggccagtagtaaaccgccta-tagacagct-c-g
VIRAL ENVELOPE PROTEIN 36 L T I T P Q A S S K R L I D S S
REVERSE TRANSCRIPTASE 1187 S N N N S P G Q ^
 Sbjct: 5679 gtctgcaa-aataagaaccccccaccagccca-t-g---ac-cctcaacttggcaggta
 env protein 43 S W Q N K N P H Q P M T L T W Q V L
 pol polyprotein 1162 S A K ^

Query: 5770 aaccccca-t--agaccc-tatcccttac-ctggctgat-tattgaccc-tgatac-g
VIRAL ENVELOPE PROTEIN 51 N P H R P L S L F W L I I D P D T
 Sbjct: 5732 -tcccaaactggagacgtgtctggatacaaaggcag-tcca-gcccccttg-gactt
 env protein 61 S Q T G D V V W D T K A V Q P P W T

Query: 5821 ggtg---tcactgtaaa-t-a-gca---c--tcgagg-t-g--ttg-cttctt-aga-
VIRAL ENVELOPE PROTEIN 68 G V T V N S T R G V A P R
 Sbjct: 5787 ggtggcccacact-taaacctgtatgtgccttggcgctagtcttgacttggat
 env protein 79 W W P T L K P D V C A L A A S L E S W D

Query: 5860 -----ggcacc--tg-gt-ggc--ctgaactgcatttctg-cctcc-gatt--gatta-
VIRAL ENVELOPE PROTEIN 81 G T W W P E L H F C L R I I
 Sbjct: 5846 atccccggaaacggatgttcgtctctaaac-g-a-gtcagacccctccggactcagactt
 env protein 99 I F G T D V S S S K R V R P P D S P Y

Query: 5903 ac-cccgtgtaaaagca---cacct----ccca-a----c--ctagtc-cg---ta-
VIRAL ENVELOPE PROTEIN 95 N P A V K S T P P N L V R
 Sbjct: 5903 acgtgcgc---tataagcaaaccctggggagccatagggtcagctaccctcggtgc
 env protein 118 T A A Y K Q I T W G A I G C S Y P R A

<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?1>

21/02/01

Query: 5942 g---t---tatgg---gttcta---tt----g-ctg-ccc---a-ggc---ac---agag
 VIRAL ENVELOPE PROTEIN 108 S Y G F Y C C P G T E
 Sbjct: 5961 ggactagaatggcaagctctacatcttacgtatgtccccggatggccggaccctttcag
 env protein 137 R T R M A S S T F Y V C P R D G R T L S

Query: 5974 aaag--agaa-atactgtggggctctgggaat-cc-t-ctgt-aggagat-gga--g
 VIRAL ENVELOPE PROTEIN 119 K E K Y C G G S G E S F C R R W S
 Sbjct: 6021 -aaactagaagggtgc-g-ggggg--ct--agaatccatactgtaaaga-atgggatg
 env protein 157 E A R R C G G L S S L Y C K E W D C

Query: 6024 ctgcgtcacctccaaacgatggagactgg-aaatggccgatct--ct-ctc-caggaccgg
 VIRAL ENVELOPE PROTEIN 136 C V T S N D G D W K W P I S L Q D R
 Sbjct: 6073 -ttagg--acc---acg--gg-gnccgggtattgg-ctatctaatactccaaangacc--
 env protein 175 E T T G T G Y W L S K S S K D

Query: 6079 gt-aaaatttcttgcattccggccoggggaaatgtc-aaaaatgtatggaaact-atat
 VIRAL ENVELOPE PROTEIN 154 V K F S P V N S G P G K Y K M M K L Y
 Sbjct: 6120 -tcataaa-c----tgtaaaat-gggacc---aaatagc-gaatgg--g--actca-a-
 env protein 190 L I T V K W D Q N S E W T Q

Query: 6136 aaagat-aagagctgtccccca-tcagacttagat-tatc----taaagata-agttc
 VIRAL ENVELOPE PROTEIN 173 K D K S C S P S D L D Y L K I S F
 Sbjct: 6161 aaatttcaacag-tg-tcaccagacccggct--ggtgtaaaccctttaaa-ataga-tttc
 env protein 204 K F Q Q C H Q T G W C N P L K I D F

Query: 6187 actgaaaaggaaaaacaggaaaa-tattcaa-aagtggataatggatgagctgggaat-
 VIRAL ENVELOPE PROTEIN 190 T E R K T G K Y S K V D K W Y E L G N
 Sbjct: 6215 ac---a---gacaa-agaaaaattatccaaggactggataacggggaaaaacctgggatc
 env protein 222 T D K G K L S K D W I T G K T W G L

Query: 6244 agttttttatratatg-g-cggggga-gcagg-gtccactttaaccattccgccttagat
 VIRAL ENVELOPE PROTEIN 209 S F L L Y G G G A G S T L T I R L R I
 Sbjct: 6269 ag----atccatgtgtctggacatccaggcgatcatacgttccatcg-cttaa-
 env protein 240 R F Y V S G H P G V Q F T I R L

Query: 6300 aqa-gacggggaca-gaaccctgtggcaatgggacccga--taaagtactggct-gaa
 VIRAL ENVELOPE PROTEIN 228 E T G T E P P V A N G P D K V L A E
 Sbjct: 6318 aaatcac--caacatg--ccagctgtggcagtagtgcctgcacct--cgtccctg-tggaa
 env protein 256 K I T N M P A V A V G P D L V L V E

Query: 6355 caggggcc-cc----cgccctgg---agccacccgataactgcc---gg--tgccc
 VIRAL ENVELOPE PROTEIN 246 Q G P P A L E P P H N L P V P
 Sbjct: 6371 caaggacccatccatgaaacgtccctgcctccaccctct--cttcccccaaggggaaacgc
 env protein 274 Q G P P R T S L A L P P P L P P R E A

Query: 6400 caattaacc-tcgtgtggccgtac-ataa---cacagcc---gc--cta--gca-acag
 VIRAL ENVELOPE PROTEIN 261 Q L T S L R P D I T Q P P S N S
 Sbjct: 6429 ca---ccggccatctc-c-cccgactctaactccacacggccctggcacttagtgcacaaac
 env protein 293 P F P S L F D S N S T A L A T S A Q T

Query: 6447 taccactggatgttttacaa---caacgcctagaa-actcc-ccagggtttctgt
 VIRAL ENVELOPE PROTEIN 277 T T G L I P T N T P R N S P G V P V
 Sbjct: 6484 tcccaec-gg--tga-gaaaaacaaatgttac-ccta-aacactccgccc---tcc--
 env protein 312 P T V R K T I V T L N T P P P P

Blast Result

```

Query: 6501 taagacaggac-agagac-tcttcagtctcatccaggagctttcc-aagccatcaact-
VIRAL ENVELOPE PROTEIN 295 K T G Q R L F S L I Q G A F Q A I N
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Sbjct: 6529 caccacagg-cgacagacttttga-tcttgatgcaggggggcttcctaa-cc-tttaatg
env protein 327 T T G D R L F D L V Q G A F L T L N

```

```

Query: 6557 ccaccggaccctgtatgcacatcttcattttgtggctttgtctat-cctca-ggg-cctccctt
VIRAL ENVELOPE PROTEIN 313 S T D P D A T S S C W L C L S S G P P
                                         | ||| | ||| | | ||||| | ||||| | ||||| | | | | | ||| | ||| | ||| |
Sbjct: 6585 ctaccaacccagggggccactgagttgtgtggctttgt-t-tggc-catggggccccccctt
env protein 345 A T N P G A T E S C W L C L A M G P P

```

Query: VIRAL ENVELOPE PROTEIN	6614 attatggggat-ccctaa--agaaaatccaatg--tga--ccaaagagcataga
	332 Y Y E G M A K E R K F N V T T K E H R
-----	-----
Sbjct: env protein	6642 attatggaaatggctcatcg---gaggtt---gcct-actccaccega-ccttg-
	364 Y Y E A I A S S G E V A Y S T D L

Query: 6723 ggggacargcataggaaaagctcccccacat-aaaaacccaaaaacatttgcata-agtactgtg
VIRAL ENVELOPE PROTEIN 369 G T C I G K A P P S H Q H I C Y S T V
 ||| ||||||| ||||| | ||| | ||||| ||| ||| ||| ||| ||| ||| |||
Sbjct: 6748 cgggttgcataagg-aaagggtcccttaccatcagcatctgtcaatcag-accc-
env protein 400 G L C I G K V P F T H Q H I C N Q T L

Query:	6950	ttggtccttgcgtga-a---	catgactatcggtataacc-	gaccaaaaagagaacccg
VIRAL ENVELOPE PROTEIN	444	V V L D E	Y D Y R Y N	R P K R E P
Subjct:	6975	t-t-ttttgcgttacaggcctatgcataat	ttctcaccccgaga-ctaaaaagagaggctg	
env protein	475	V L Q A Y D N	S H P R T K R E A	

Query: 7113 acatggggccat-gacagaagatccggag-aatttaaggagtctgttagccatcgaa
21/02/01
<http://www.ncbi.nlm.nih.gov/blast/b1/2seq/wblast2.cgi?1>

Blast Result

VIRAL ENVELOPE PROTEIN 499 H A A M T E D L R A L K E S V S N L E
 Sbjct: 7140 agat-c-gccataga-tgcgtgactccggccctccaa-gactcagtcagaaggtagag
 env protein 530 Q I A I D A D L R A L Q D S V S K L E

Query: 7171 gagtcctgacttctttgtctgaagtggccctacagaaccggag-gggatctaga-tctgc
 VIRAL ENVELOPE PROTEIN 518 E S L T S L S E V V L Q N R R G L D L
 Sbjct: 7196 gactcaactgacttccctgtccgaggttagtgcctccaaataggagaggcctt-gact-tgc
 env protein 549 D S L T S L S E V V L Q N R R G L D L

Query: 7229 tttttctaaagagaagggtgggttatgtgcagccctaaaagaagaatgttgtatgttag
 VIRAL ENVELOPE PROTEIN 537 L F L R E G G L C A A L K E E C C F Y V
 Sbjct: 7254 tttttctaaaagaagggtggccctgtgcggccctaaaggaaagagtgtgttttacata
 env protein 568 I F L K E G G L C A A L K E E C C F Y I

Query: 7289 atcaactcaggagcca-t-cagagactccatgaacaagct--tagaaaaa--agtttagaga
 VIRAL ENVELOPE PROTEIN 557 D H S G A Y R D S M N K L R K K L E
 Sbjct: 7314 accactcaggtg-cgtac-gggactccataaaaactaaaggaaaactggata-aaa
 env protein 588 D H S G A V R D S M K K L K E K L D K

Query: 7343 ggc-g-tcgaaggaaaagagaggctgaccagggtggcttgaaggatggcaacaggc
 VIRAL ENVELOPE PROTEIN 575 R R R R R R E A D Q G W F R G W F N R S
 Sbjct: 7371 gacagtttag-agcccaaaaaag---ccaaaactggatgtcaaggatggcaataactc
 env protein 607 R Q L E R Q K S Q N W Y E G W F N N S

Query: 7401 tcctggatgaccacccctgttttgc--tctgacggggcccttagtagtctgtctctg
 VIRAL ENVELOPE PROTEIN 595 P W M T T L L S A L T G P L V V L L L
 Sbjct: 7426 cccttggttcaactaccctgtatcaaccatc-g-ctggggcccttattacttcttctg
 env protein 626 P W F T T L L S T I A G P L L L L L

Query: 7459 ttacttacagt--t-gggcctgtttaattataggttgtgc-ctttggtaga-ga-a
 VIRAL ENVELOPE PROTEIN 614 L L T V G P C L I N R F V A F V R E
 Sbjct: 7484 ttgtc--ca-ccctggccatgcatcatcaataaggtagt-caattcatca-atgata
 env protein 645 L L I L G P C I I N K L V Q F I N D

Query: 7513 cgaggatgtca-gtccaga-tcatggactttaggcaacagtaccaaggcctt-gagc
 VIRAL ENVELOPE PROTEIN 632 R V S A V Q I M V L R Q Q Y Q G L L S
 Sbjct: 7539 gga-taagtgcgtttaaaaattc-tggtctttagaca-atacc-aggcc--ctagag-
 env protein 663 R I S A C ^

Query: 7570 caaggagaaactgacccttagccctccagttctaaggatagaacttataacaagacaag
 VIRAL ENVELOPE PROTEIN 651 Q G E T D L ^
 Sbjct: 7592 -aa--cgaggtaacottaattt---gttataaggatgtatcacaaagaga-

Query: 7630 aagtggaaatgaaagg-a-tg--aaaatgcaacctaaccctccag-aac-ccaggaagt
 Sbjct: 7644 atg-gggaaatgaaaggatgtttttttagcc-aa--ct-gcagtaacgcc---att

Query: 7685 taata--aaaagcttaatgcggcccg--aattacagaccct--g--ctggctggcagta
 Sbjct: 7696 tgcgtggcaca-c-ctaaa-ggataggaaaaatacag---ctaagaacagg--gcc---a

Query: 7737 aatagg-ta---gaaggta--cacttccattgttccaggccctgtatctggctaa
 Sbjct: 7745 aacaggatatctg-tggtcatgcac--c---tggcc---ccggc---ccaggcc-aa

Query: 7791 gta-aga---taacaggaaatgagttgact-aatc-gcttat=tggatt-ctgtaa-aa
 http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?1 21/02/01

Blast Results

```

Sbjct: 7790 ggacagagggttccca-gaaat-agatgagtcaa-cagc--a---gtttccagcaaggaa

Query: 7842 ctgactgg--caccataga-agaattga-ttacacatggacag--cccta--gtgaccta
Sbjct: 7841 caga-gggttc-ccagaaataga--tgagtca-acs--g-cagtttcc-agggtgcc--t

Query: 7894 tctcaactgcaatctgtcaactctgcccaggagccacgcagatgcggaccccgagcta
Sbjct: 7890 -ctcaacccg---t-t-tca-----agga-ctc-c-c--at---ga---ccgg-g---t

Query: 7954 ttttaaaatgattggtccacggagcgcgggtctcgatattttaaaatgattggtcc-at
Sbjct: 7921 -----aat-t-----cac----ccctggc-ct---ta-tttgaactaa----ccaaat

Query: 8013 ggagcgc-gggctctcgatattttaaaatgattggttgtgacgcacag-gctttgttg-
Sbjct: 7954 -ta-c-cttgcctctcgct-----tctgt-ac-c-c-gcgcttt-ttgc

Query: 8070 tga-accccataa-aagctgtcccgattc-cgcactcgggc-cgcagtccttacccct
Sbjct: 7990 t-ata-----a-ataaactt---cagaaactc-cacccggagcgc-cagtcct-ta-----t

Query: 8126 gcgtg-gtgcgttgc---cgcactgtggggcccccagggcgttggaaataaaaaatcttgcgttt
Sbjct: 8033 gagagactg-agccggcccg-ggttaccc-gtgtg--tccaataaaa--cctcttgcgtgatt

Query: 8183 gca 8185
Sbjct: 8086 gca 8088

```

Blast Result

```

Query: 7635 gg-gaa-----tgaag-ga--tgaat-gc-aacca-ac----c---ctc---
Sbjct: 352   |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||

Query: 7669 ccag--aaccca-gga-a-gt--ta--at-a---a--a--a---a-gctcaaatg--c
Sbjct: 410   cctggtaatcaggagacgtcccaggatcagggaggaccaggacgc-ct-ggtggac

Query: 7705 ccc-cg--aa---t--tacagacccctgtggc-t-gccagtaaataggt-agaag-g-t
Sbjct: 468   |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||

Query: 7751 cacactccct-attgt-te-c-agg--gact---gatac-ctgg---c-ctaag-taa
Sbjct: 524   |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||

Query: 7795 gataacaggaaa---tgat-tgactaatac--gt-t---atc-tggat-t---c-t
Sbjct: 582   g-tttctggaaaggccctgtgtcac-agtcttcttccaaatcatcaatatggacaa
1                               M G Q

Query: 7836 g-taa---a--actga-ct--ggca-c---at-a--gaaga-at-tgattacaca---
Sbjct: 640   gataattctaccctt-actccctcaactctaataactggagagatgtgagaaca-aggg
4       D N S T P   I S L T L N H W R D V R T R

Query: 7874 ttgacagccota---gt-gacctatctcaactgcaatctgtc--ac--tctg--cc-c-
Sbjct: 698   ctcaca-atotatccgtggaa---atccaaaaaggaaaaatggcagacttctgttccca
23      X H N L S V   I K K G K W Q T F C S S

Query: 7921 ag-gagcccac----gcagatgcg--gacccggag---ct-----
Sbjct: 755   agtg-qccccatccggc-g-tgggtggccacccggagggaaacttttaatctctgtca
42      E W P T F G   V G W P P B G T F N L S V

Query: 7954 tt-----ttaaatgattggtc--cacggag--cgccgg-c-ttc-ga--tatttt
Sbjct: 812   ttttgcgttaaaaagatt-gtccttca-ggagaacggggacatc-cgaccaggatc
61      I F X V K K I   V P Q E N G G H P D Q V

Query: 7997 aaaatgtat--tggccatgg-ag--cgcgggttc-gatat--tttaaatgattgg-
Sbjct: 869   catat-atcggtt--atggcaggac-ctcgc-c-caga-atccccccatcg---ggtg
80      P Y I V V W Q D L A Q N P P P W V

Query: 8048 -tcg--tgacg-cacag---gtt-ttgt-tggat-acc-----ccat-----a-
Sbjct: 919   ccagct-cgcaca-aggtcgatgtgtctctgataccggaaagaccaggatgcggggaggc
97      P A S A K V A V V S D T R R P V A G R

Query: 8081 -a-aatgt-gt-cgcattccgca-ct---cg---ggccg-c---agtcct-ct-----
Sbjct: 977   catcgatctcccccggccccatctaccggaaacagacgacttactctctctgt
116      P S A P P R P I Y P A T D D L L L S

Query: 8120 -accc-tcg---cg---t---gg-tgtac-g---a---ctgtggcc-cagccgcct
Sbjct: 1037 aacccac-gccccccgcctatccggcactgcacccctctggccctcgatgcgt
136      E P T P P P Y P A A L P P P L A P Q A I

Query: 8158 tggat-----ataa-----aaat-cctttgtct-gttgc-atac 8187

```

<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?1>

21/02/01

Blast Result

Subject: 1095 cggaccggcgtcaggccagatgcc-c--gatag-tagcgate 1132
gg polyprotein 156 G p p S G Q M P D S S D
CPU time: 6.79 user secs. 0.19 sys. secs. 6.98 total secs.

Gapped
Lambda X H
0.0514 0.0120 0.0135

Gapped
Lambda X H
0.0514 0.0120 0.0135

Matrix: blastn matrix:5 -2
Gap Penalties: Existence: 1, Extension: 2
Number of Hits to DB: 7964
Number of Sequences: 0
Number of extensions: 7964
Number of successful extensions: 13
Number of sequences better than 10.0: 1
length of query: 8209
length of database: 2,812,937,656
effective HSP length: 1920
effective length of query: 6289
effective length of database: 2,812,935,736
effective search space: 17690552843704
effective search space used: 17690552843704
T: 0
A: 0
X1: 135 (10.0 bits)
X2: 1347 (99.9 bits)
S1: 250 (24.9 bits)
S2: 463 (40.7 bits)

<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?1>

21/02/01

Biast Resut

Attachment B



Blast 2 Sequences results

PubMed

Entrez

BLAST

0MM

Taxonomy

Structure

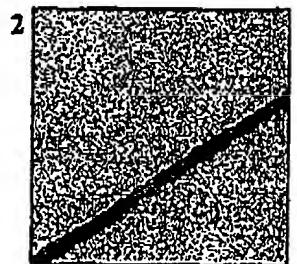
BLAST 2 SEQUENCES RESULTS VERSION BLASTN 2.1.2 [Nov-13-2000]

Match: Mismatch: gap open: gap extension:

x dropoff: expect: wordsize: Filter

Sequence 1 gi:4538106 Sequence 3 from Patent WO9740167. Length 1575 (588 .. 2162)

Sequence 2 gi [9630311](#) Gibbon leukemia virus, complete genome | Length 2370 (631 .. 3000)



NOTE: The statistics (bitscore and expect value) is calculated based on the size of nr database

NOTE: If protein translation is reversed, please repeat the search with reverse strand of the query sequence

Score = 362 bits (4802), Expect = 2e-98
Identities = 1164/1749 (66%), Gaps = 345/1749 (19%)
Strand = Plus / Plus

Query: 588 atggggac-ag2cagtgactaccccccctagt-t---tgactctcgaccattggact
 VIRAL CORE PROTEIN 1 M G Q T V T T P L S L T L D H W T
 Sbjct: 651 atgggacaagataat-tctacccc---ta-tctcccccactctaaatcactggaa--
 gag polyprotein M G Q D N S T P I S L T L N H W

Query: VIRAL CORE PROTEIN 38	698	tttctgtgcctctgtatggccaaacattcgatgttggatggccatcagaggggac
		F C A S E W P T F D V G W P S E G T
Sbjct: gag polyprotein	741	tttctgtttctcccgactggccaaacattcggtgtgggtggccaccggagggact
		F C S S E W P T F G V G W P P E G T

Query: 758 t---tctgaaattatccggctgttaaggcaattt---tttca-gactggacc
 VIRAL CORE PROTEIN 57 S E I I L A V K A I I F Q T G P
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Sbjct: 801 tctctctgtcatt-t--trgcagttaa---azagattgtcttcagga---ga-ac
<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?0> 22/02/01

Blast Result

gag polyprotein	57	L S V I F A V K K I V F Q E N
-----------------	----	-------------------------------

Query: 811	ctcatccatgtcaggag--ccctatatacct--tacgtggcaagatgg--cagaa
VIRAL CORE PROTEIN 75	S H P D Q E P Y I L T W Q D L A E

Sbjct: 851	gacatccggacca--agtccatatacgtgta--tggcaggaccccccaga-
gag polyprotein 74	G H P D Q V P Y I V V W Q D L A Q

Query: 865	ctccgcacatgggttaaaccatggctaaataaaccggaaagcc-agggtcccccaat
VIRAL CORE PROTEIN 93	P P P W V K P W L N K P R K P G P R I

Sbjct: 905	ccccacccatgggt--gcac-gcct-----cc-----gcacaggcgctg--t
gag polyprotein 92	P P P W V P A S A K V A V

Query: 924	gctcttggaga---gaaaaacaaacact--c-----ggccgaaaaagtcgagc-cc
VIRAL CORE PROTEIN 113	A L G E K N K H S A E K V E P

Sbjct: 944	tctct--gatacccg-aagac---cagttggggggaggcc-a----tc-agctcc
gag polyprotein 105	V S D T R R P V A G R P S A P

Query: 973	cctcg-----tatotaccccgag-atac-gaggagc---cgccgact-tggccgg
VIRAL CORE PROTEIN 129	S S Y L P R D R G A A D L A G

Sbjct: 989	cc-cgaccccccacata-cccg-gcaacagacga-cttactctctct---ctga
gag polyprotein 120	P R P P I Y P A T D D L L L S E

Query: 1021	caacctgtccccccaccccccattatccagcacagggtgctgtgaggggacccct-gc
VIRAL CORE PROTEIN 145	P T C S P T P L S S T G C C E G T S A

Sbjct: 1042	--ac--g--cccccggcc---tateccggc---ggactgcca--cccaactctggc
gag polyprotein 138	T P P P Y P A A L P P P L A

Query: 1080	c---c--t-gga--gc--tccgg----tg---g-t-ggaggg-a-cctg-----c
VIRAL CORE PROTEIN 165	P G A P V V E G P A

Sbjct: 1087	caggcgatccggacccccgtcaaggccagatgcccgtatgtccgatccgtggggcc
gag polyprotein 153	Q A I G P P S G Q X P D S S D P E G P

Query: 1113	gggactcgagccggagaggcgcc-accccgag--cgacacagacgaga-tcgcg
VIRAL CORE PROTEIN 176	G T R S R R G A T P E R T D E I A

Sbjct: 1147	g---ctggga-ccag-gagtccgcgtgcccgcagtccag-cagac-a-actcg-gg
gag polyprotein 173	A G T R S R R A R S P A D N S G

Query: 1169	acc-g-ctgcgcacct---a---tggccctc---ccat-g---ccagg--g-gg
VIRAL CORE PROTEIN 194	P L R T Y G P P M P G G

Sbjct: 1195	-cctgact-c-ca-ctgtgattitccgcggccataggaccccccggccgagc
gag polyprotein 189	P D S T V I L P R A I G P P A E

Query: 1209	ttgc--agccctc--cagtattggccctttcttcgtcagatctctataattgga
VIRAL CORE PROTEIN 208	L Q P L Q Y W P F S S A D L Y N W

Sbjct: 1251	tggcctggcccttacaatattggccctttcttcgtcagatcttataattgga
gag polyprotein 208	G L V P L Q Y W P F S S A D L Y N W

Query: 1265	taaccatecccccttctcgaggatcccc--acgcctcacgggg-t--tggtgga
VIRAL CORE PROTEIN 227	N H P P F S E D P Q R L T G L V E

http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?0 22/02/01

Blast Result

Attachment C

 NCBI

Blast 2 Sequences results

PubMed

Entrez

BLAST

OMIM

Taxonomy

Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTN 2.1.2 [Nov-13-2000]

Match: Mismatch: gap open: gap extension:

Sequence 2 gi 9630311 Gibbon leukemia virus, complete genome Length 6056 (1..6056)

21

NOTE: The statistics (bitscore and expect value) is calculated based on the size of nr database.

NOTE: If protein translation is reversed, please repeat the search with reverse strand of the query sequence

Score = 798 bits (10671), Expect = 0.0
Identities = 2598/3890 (66%), Gaps = 667/3890 (17%)
Strand = Plus / Plus

Query: REVERSE TRANSCRIPTASE	2331 -cagttatccatggactacccgaagaaccgtt---gacttggga-gtggggacg
	57 Q Y P W T T R R T V D L G V G R
Sbjct: pol polyprotein	2363 tc---tacccctggaccacaaaaaga--cttttaaaaatt-ggacat-aaaca
	57 V Y P W T T K R L L K I G H K Q

Query: 2386 actcgttctggcat-ccctgagtgcccaag-taccctt-tcttaggttagaga
 REVERSE TRANSCRIPTASE 75 H S F L V I P E C P V P L L G R D
 Sbjct: 2417 actcccttcctggcataccg-gagtgcctgtct-cctctgt-tgggcaggaa
 pol polyprotein 75 R S F L V I P E C P A P L L G R D
<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?0> 22/02/01

Blast Result

```

Query: 2443 ccaaga-t---gggagctcaaatttc---ttrr--gaac-aaggaaagaccag
REVERSE TRANSCRIPTASE 94 r K H G A Q I S F E Q G R P
Sbjct: 2474 ccaa-actaaag---qccccaga--tccagtttccg--ctgagg--gccca-
pol polyprotein 94 T K L K A Q I Q F S A E G P

Query: 2491 c-tgtgaataacaaacccatcaactgtgttacc--ctcc-aa--tttagatga
REVERSE TRANSCRIPTASE 110 S V N N K P I T V L T I Q L D P
Sbjct: 2522 catggggagaac--gccct-actatg-tg-cctggctctaaaccttggaaaga
pol polyprotein 110 T W G E R P T M C L V L N L E E

Query: 2545 gactatattctcccaagtaaagcctg-a---tcaagatat--a--cagtcc
REVERSE TRANSCRIPTASE 128 R L Y S P Q V K P D Q D I Q S
Sbjct: 2576 gacta-----catgaaaagccatgaccctc-ctctatcgttccaa-ccca-tcc
pol polyprotein 128 R L H E K P V P S S I D P S

Query: 2597 gcagtttccca-agcctggcagaaaccgcag--gg-atgggtttggcaaa
REVERSE TRANSCRIPTASE 146 Q F P Q A W A E T A G M G L A K
Sbjct: 2625 gc-ttttcccactgtatggcagaaa--g-aqccggatggacttagccaa
pol polyprotein 145 L F P T V W A E R A G M G L A N

Query: 2653 ccccacag-gttattca-a-ctgaaggcca-gtg-ctacaccagtatcag-t
REVERSE TRANSCRIPTASE 164 P P Q V I Q L K A S A T P V S V
Sbjct: 2680 ccac-cagtggtagt-agagct-aagatcaggtgcct-caccaagtggctgtt
pol polyprotein 163 P P V V V E L R S G A S P V A V

Query: 2707 accccttgaggtagagaggctcgagaaggaatttggccgatgttca-aagat
REVERSE TRANSCRIPTASE 182 Y P L S R E A R E G I W P H V Q R
Sbjct: 2735 atccaatgagcaaagactggaaaggatcagacccacatcccaaag-t
pol polyprotein 181 Y P M S K E A R E G I R P H I Q K

Query: 2765 -ac--aggccatcttagt-tcttattcaatc-cccttggaaatactccc-ctg
REVERSE TRANSCRIPTASE 202 Q G I L V P V Q S P W N T P L
Sbjct: 2791 gacctaggg-gtcttggccctgt-cggtcgccc-tggaaatac-cccttctg
pol polyprotein 200 D L G V L V P C R S P W N T P L

Query: 2819 tggaaaggctgggaccaatgattatcgaccagtacaggacttgagagaggtc
REVERSE TRANSCRIPTASE 220 R K P G T N D Y R P V Q D L R E V
Sbjct: 2847 aaaaaagccagggaccaatgactatggcoagttcaagacacttgagagaaaatt
pol polyprotein . 219 K K P G T N D Y R P V Q D L R E I

Query: 2879 ggtgcaggacataca-cccaacggtcccgaaaccttataa-cctcttggcg
REVERSE TRANSCRIPTASE 240 V Q D I H P T V P N P Y N I L S
Sbjct: 2907 ggtacaggatattoatccc-acagtcctaaaccttacaatcttc-tgatgtt
pol polyprotein 239 V Q D I H P T V P N P Y N L L S

Query: 2937 cctgaacgga-ac-tgttacacagtatttgg-a-ctaaaagatgctttttctt
REVERSE TRANSCRIPTASE 259 P E R N W Y T V L D L K D A F F
Sbjct: 2965 cct-aga-tatacttctgtactcagtcttagatc-tcaaggatgcctttttctt
pol polyprotein 259 P E R N W Y T V L D L K D A F F

```

168 23 2001 05:54

0141 221 7739

PAGE . 33

Blast Result

Page 1 of 5

Attachment ▷



Blast 2 Sequences results

PubMed

Entrez

BLAST

OMIM

Taxonomy

Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTN 2.1.2 [Nov-13-2000]

Match: Mismatch: gap open: gap extension:
 x dropoff: expect: wordsize: Filter

Sequence 1 gi_4538106 Sequence 3 from Patent WO9740167. Length 1971 (5620.. 7590)

Sequence 2 gi_9630311 Gibbon leukemia virus, complete genome Length 2449 (5552.. 8000)

2 1

NOTE: The statistics (bitscore and expect value) is calculated based on the size of nr database

NOTE: If protein translation is reversed, please repeat the search with reverse strand of the query sequence

Score = 410 bits (5448), Expect = e-112
 Identities = 1396/2290 (60%), Gaps = 560/2290 (24%)
 Strand = Plus / Plus

Query: 5620 atgcatacccacgttaagcggggccacccccc-actc-g--ggg---tggaa
 VIRAL ENVELOPE PROTEIN 1 M H P T L S R R H L P T R G G
 REVERSE TRANSCRIPTASE 1153 R A S H V K P A P P P D S G W
 ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Sbjct: 5573 atgcttcteacaccaaacctgcaccacccatcgcac-cagatgagtctgg-
 env protein 8 M L L T S N L X H L F H Q M S P G
 pol polyprotein 1127 H A S H L K P A P P S A P D E S W

Query: 5672 aaagactgagaatccccttaagcttcgcctcca-tc-gc--ctggttcctta
 VIRAL ENVELOPE PROTEIN 18 K R L R I P L S F A S I A W F L
 REVERSE TRANSCRIPTASE 1171 K T E N P L K L R L H R L V P Y
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Sbjct: 5631 aaagactgatcatcctttaagct--gcgt--attcggggc-gg--cggga
 env protein 27 K R L I I L S C V F G G G G
 pol polyprotein 1147 K T D H P L K L R I R R R R D

Query: 5724 -aacaaat---aactcc--ccag----g-c---ca-----gta--gt--
 VIRAL ENVELOPE PROTEIN 36 T I T P Q A S S
 REVERSE TRANSCRIPTASE 1188 N N N S P G Q ^ ^
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Sbjct: 5684 caa-aataagaaccccccacccatgaccctcaacttggcaggatacttcc
 env protein 45 Q N K N P H Q P M T L T W Q V L S
 pol polyprotein 1164 A K ^ ^

Query: 5754 ---ccat-----at----agacacgtcgaaaccccatagacatt--tato
 VIRAL ENVELOPE PROTEIN 46 L I D S S N P H R P L S
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Sbjct: 5743 agacgttgcgtggatacaaggcag-tccagcccccattggc-cttgggtggc
 env protein 65 D V V W D T K A V Q P P W T W N

Blast Result

Page 2 of 5

Query: 5794 --acctggotgattat-tgaccctgatacgggt-gtcaactgtaaatagcact
 VIRAL ENVELOPE PROTEIN 59 T W L I I D P D T G V T V N S T
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Sbjct: 5801 aaacct-gatg--tatgtg-ccttg--gcccgtatgt--ct-tgagt--c-ct
 env protein 84 K P D V C A L A A S L E S

Query: 5649 ttgcctcttagaggcacc--tggt--ggc--ctgaactgcattctg-cctc
 VIRAL ENVELOPE PROTEIN 77 V A P R G T W W P E L H F C L
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Sbjct: 5847 t--c-cc--g-gaaaccat-gtctcgctctctaaac-g-a-gtcaagaccc
 env protein 99 I P G T D V S S S K R V R P

Query: 5893 gatta-ac-cccgctgttaaaagca---cacct----ccca-a----c--
 VIRAL ENVELOPE PROTEIN 94 I N P A V K S T P F N
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Sbjct: 5897 gactatactgcgc--ttataagcaaatcacctggggagccatagggtgcag
 env protein 116 D Y T A A Y K Q I T W G A I G C S

Query: 5939 g---tag---t---tatgg---gttcta--ttgct--g----ccc---a-g
 VIRAL ENVELOPE PROTEIN 107 R S Y G F Y C C P
 ||||| ||||| ||||| ||||| ||||| |||||
 Sbjct: 5955 gggctaggactagaatgcaagctctacctt-ctacgtatgtccccggatg
 env protein 135 R A R T R M A S S T F Y V C P R D

Query: 5969 ----cagagazag--azaa-atactgtgggggtotggggaat-cct-t-at
 VIRAL ENVELOPE PROTEIN 117 T E K E K Y C G G S G E S F
 ||||| ||||| ||||| ||||| |||||
 Sbjct: 6014 ctttcag--aagctagaagggtgc-g-ggggg--ct--agaatccctatact
 env protein 155 L S E A R R C G G L E S L Y

Query: 6018 at-gga--gctgcgtcacctccaaacgatggagactgg-aaatggccgatc--
 VIRAL ENVELOPE PROTEIN 134 W S C V T S N D G D W K W P I
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Sbjct: 6064 atgggattg-tgaq--acc---acg--gg-gaccggatttgg-ctatcta
 env protein 172 N D C E T T G T G Y W , L S

Query: 6070 -caggacgggt-aaaattctcccttgcataattccggccggcaagtacaa
 VIRAL ENVELOPE PROTEIN 151 Q D R V K F S F V N S G P G K Y K
 ||||| ||||| ||||| |||||
 Sbjct: 6112 aaaagacc--tcataaa-c-----tqtaaaat-----ggg-a---ccaa
 env protein 187 K D L I T V K W D .Q

Query: 6124 atgaaact-atataaaagat---aagagatgtcccca-tcaagacttagat-t
 VIRAL ENVELOPE PROTEIN 169 M K L Y K D X S C S P S D L D
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Sbjct: 6151 atg-gactca---aaa-atttcaacag-tg-tcaccagacccgtt--gggt
 env protein 201 W T Q K F Q Q C H Q T G W C

Query: 6173 taaagata-agtttactgaaaggaaacaggaaaa-tatcaa-aagtgg
 VIRAL ENVELOPE PROTEIN 185 L K I S F T E R K T G K Y S K V D
 ||||| ||||| ||||| |||||
 Sbjct: 6202 taaa-ataga-tttcac--a---gacaa-aggaaaattatccaaaggactgga
 env protein 218 K I D F T D K G K L S K D W

Query: 6230 atgagctggggat-agtttttattatatg-g-cggggga-gcagg-gtcc
 VIRAL ENVELOPE PROTEIN 204 Y E L G N S F L L Y G G G A G S
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Sbjct: 6254 aaaacctggggatataag-----attctatgtgtctggacatccaggctac
 env protein 235 K T W G L R F Y V S G H P G V

<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?0>

22/02/01